

## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

- (i) APPLICANT: POWELL, Jerry S.
- (ii) TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL  
EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Dorsey & Whitney, LLP
  - (B) STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
  - (C) CITY: Seattle
  - (D) STATE: Washington
  - (E) COUNTRY: USA
  - (F) ZIP: 98101
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: MS-WORD
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: Not Yet Assigned
  - (B) FILING DATE: Concurrently Herewith
- (viii) PREVIOUS APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/466,412
  - (B) FILING DATE: June 6, 1995
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Roberts, Mark W.
  - (B) REGISTRATION NUMBER: 46,160
  - (C) REFERENCE/DOCKET NUMBER: 500582.12
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 206-903-8748
  - (B) TELEFAX: 206-903-8820

### (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2426 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCCGGTGTGG TCACCCGGCG CGCCCCAGGT CGCTGAGGGA CCCC GGCCAG  
GCGCGGAGAT 60

GGGGGTGCAC GGTGAGTACT CGCGGGCTGG GCGCTCCCGC CCGCCCGGGT  
CCCTGTTTGA 120

GCGGGGATTT AGCGCCCCGG CTATTGGCCA GGAGGTGGCT GGGTTCAAGG  
ACCGGCGACT 180

TGTCAAGGAC CCCGGAAGGG GGAGGGGGGT GGGGCAGCCT CCACGTGCCA  
GCGGGGACTT 240

GGGGGAGTCC TTGGGGATGG CAAAAACCTG ACCTGTGAAG GGGACACAGT  
TTGGGGGTTG 300

AGGGGAAGAA GGTTTGGGGG GTTCTGCTGT GCCAGTGGAG AGGAAGCTGA  
TAAGCTGATA 360

ACCTGGGCGC TGGAGCCACC ACTTATCTGC CAGAGGGGAA GCCTCTGTCA  
CACCAGGATT 420

GAAGTTTGGC CGGAGAAGTG GATGCTGGTA GCCTGGGGGT GGGGTGTGCA  
CACGGCAGCA 480

GGATTGAATG AAGGCCAGGG AGGCAGCACC TGAGTGCTTG CATGGTTGGG  
GACAGGAAGG 540

ACGAGCTGGG GCAGAGACGT GGGGATGAAG GAAGCTGTCC TTCCACAGCC  
ACCCTTCTCC 600

CTCCCCGCCT GACTCTCAGC CTGGCTATCT GTTCTAGAAT GTCCTGCCTG  
GCTGTGGCTT 660

CTCCTGTCCC TGCTGTCGCT CCCTCTGGGC CTCCCAGTCC TGGGCGCCCC  
ACCACGCCTC 720

ATCTGTGACA GCCGAGTCCT GCAGAGGTAC CTCTTGGAGG CCAAGAAGGC  
CGAGAATATC 780

ACGGTGAGAC CCCTTCCCCA GCACATTCCA CAGAACTCAC GCTCAGGGCT  
TCAGGGA ACT 840

CCTCCCAGAT CCAGGAACCT GGCAC TTGGT TTGGGGTGGG GTTGGGAAGC  
TAGACACTGC 900

CCCCCTACAT AAGAATAAGT CTGGTGGCCC CAAACCATAC CTGGAAACTA  
GGCAAGGAGC 960

AAAGCCAGCA GATCCTACGC CTGTGGCCAG GGCCAGAGCC TTCAGGGACC  
CTTGACTCCC 1020

CGGGCTGTGT GCATTTCAGA CGGGCTGTGC TGAACACTGC AGCTTGAATG  
AGAATATCAC 1080

TGTCCCAGAC ACCAAAGTTA ATTTCTATGC CTGGAAGAGG ATGGAGGTGA  
GTTCTTTTT 1140

TTTTTTTTTT CCTTTCTTTT GGAGAATCTC ATTTGCGAGC CTGATTTTGG  
ATGAAAGGGA 1200

GAATGATCGA GGGAAAGGTA AAATGGAGCA GCAGAGATGA GGCTGCCTGG  
GCGCAGAGGC 1260

TCACGTCTAT AATCCCAGGC TGAGATGGCC GAGATGGGAG AATTGCTTGA  
GCCCCGGAGT 1320

TTCAGACCAA CCTAGGCAGC ATAGTGAGAT CCCCCATCTC TACAAACATT  
TAAAAAATT 1380

AGTCAGGTGA AGTGGTGCAT GGTGGTAGTC CCAGATATTT GGAAGGCTGA  
GGCGGGAGGA 1440

TCGCTGGAGC CCAGGAATTT GAGGCTGCAG TGAGCTGTGA TCACACCACT  
GAACTCCAGC 1500

CTCAGTGACA GAGTGAGGCC CTGTCTCAAA AAAGAAAAGA AAAAAGAAAA  
ATAATGAGGG 1560

CTGTATGGAA TACGTTTCATT ATTCATTCAC TCACTCACTC ACTCATTTCAT  
TCATTTCATTC 1620

ATTCAACAAG TCTTATTGCA TACCTTCTGT TTGCTCAGCT TGGTGCTTGG  
GGCTGCTGAG 1680

GGGCAGGAGG GAGAGGGTGA CATCCCTCAG CTGACTCCCA GAGTCCACTC  
CCTGTAGGTC 1740

GGGCAGCAGG CCGTAGAAGT CTGGCAGGGC CTGGCCCTGC TGTCGGAAGC  
TGTCCTGCGG 1800

GGCCAGGCCC TGTTGGTGAA CTCTTCCCAG CCGTGGGAGC CCCTGCAGCT  
GCATGTGGAT 1860

AAAGCCGTCA GTGGCCTTCG CAGCCTCACC ACTCTGCTTC GGGCTCTGGG  
AGCCCAGGTG 1920

AGTAGGAGCG GACACTTCTG CTTGCCCTTT CTGTAAGAAG GGGAGAAGGG  
TCTTGCTAAG 1980

GAGTACAGGA ACTGTCCGTA TTCCTTCCCT TTCTGTGGCA CTGCAGCGAC  
CTCCTGTTTC 2040

CTCCTTGGCA GAAGGAAGCC ATCTCCCCTC CAGATGCGGC CTCAGCTGCT  
CCTCTCCGAA 2100

CAATCACTGC TGACACTTTC CGCAAACCTCT TCCGAGTCTA CTCCAATTTC  
CTCCGGGGAA 2160

AGCTGAAGCT GTACACAGGG GAGGCCTGCA GGACAGGGGA CAGATGACCA  
GGTGTGTCCA 2220

CCTGGGCATA TCCACCACCT CCCTCACCAA CATTGCTTGT GCCACACCCT  
CCCCCGCCAC 2280

TCCTGAACCC CGTCGAGGGG CTCTCAGCTC AGCGCCAGCC TGTCCCATGG  
AACTCCAGT 2340

GCCACCAATG ACATCTCAGG GGCCAGAGGA ACTGTCCAGA GAGCAAACCTCT  
GAGATCTAAG 2400

GATGTCACAG GGCCAACTTG AGGGCC

2426

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala Pro Xaa Arg Leu Ile Leu Asp Ser Arg Val Leu Glu Arg Tyr Leu  
1            5            10            15

Leu Glu Ala Lys Glu Ala Glu Xaa Ile Thr Asp Gly Gly Ala  
20            25            30

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Glu Ala Lys Glu Ala Glu Asn  
1            5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTYTCDGCGYT CYTTDGCTTC

20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Glu Asn Ile Thr Asp Gly

1            5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGCTCCTCCA TCAGTATTAT TTY

23

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.